A Gibbs' Sampler for the Parameters of a Truncated Multivariate Normal Distribution

William Griffiths

Department of Economics

University of Melbourne

Abstract

The inverse distribution function method for drawing randomly from normal and truncated normal distributions is used to set up a Gibbs' sampler for the posterior density function of the parameters of a truncated multivariate normal distribution. The sampler is applied to shire level rainfall for five shires in Western Australia.

1. Introduction

The truncated multivariate normal distribution is a reasonable distribution for modelling many natural occurring random outcomes. An example, and the one pursued in this paper, is the distribution of rainfalls in adjacent geographical areas. The usefulness of modelling rainfall in this way, and the way in which it contributes to wheat yield uncertainty, is illustrated in Griffiths et al (2001). Another area where truncated normal distributions have been used is in the modelling of firm efficiencies through stochastic production frontiers. For a general review see Greene (1997), and for one with a Bayesian flavour see Koop and Steel (2001). Posterior inferences about the location vector and scale matrix for a truncated multivariate normal distribution are complicated by the presence of a multivariate normal integral that depends on these unknown parameters. We show how to solve this problem by using latent variables that are corresponding non-truncated multivariate normal random variables; a relatively simple Gibbs' sampler involving only draws from conditional non-truncated normal distributions is set up.

The plan of the paper is as follows. Some notation and preliminaries are established in Section 2. Section 3 contains a description of the Gibbs' sampler for a truncated univariate normal distribution. This algorithm is generalised to a multivariate distribution in Section 4. An application to shire-level rainfall is given in Section 5.

2. Notation and Preliminaries

Let y be an $(N \times 1)$ normal random vector with mean vector μ and covariance matrix Σ. Its pdf is given by

$$f(y \mid \mu, \Sigma) = (2\pi)^{-N/2} \mid \Sigma \mid^{-1/2} \exp\left\{ -\frac{1}{2}(y-\mu)' \Sigma^{-1} (y-\mu) \right\} \tag{1}$$

Suppose x is a truncated version of y; it has the same location and scale parameters μ and Σ, but is truncated to the region $R = \{(a_i < x_i < b_i),\ i = 1,2,\ldots N\}$. We include cases where some or all of the a_i could be $-\infty$ and some or all of the b_i could be $+\infty$. The pdf for x is given by

$$f(x \mid \mu, \Sigma) = \left[P(\mu,\Sigma)\right]^{-1} (2\pi)^{-N/2} \mid \Sigma \mid^{-1/2} \exp\left\{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right\} I_R(x) \tag{2}$$

where $I_R(x)$ is an indicator function equal to one when x is in the region R and zero otherwise, and $\left[P(\mu,\Sigma)\right]^{-1}$ is a modification to the normalising constant, attributable to the truncation. Specifically,

$$P(\mu,\Sigma) = \int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_N}^{b_N} f(y \mid \mu, \Sigma)\, dy \tag{3}$$

Let $\boldsymbol{x} = (x_1, x_2, ..., x_T)$ denote a random sample from the truncated multivariate normal pdf $f(x \mid \mu, \Sigma)$. The pdf for this random sample $\boldsymbol{x}$ (likelihood function for μ and Σ) is

$$\begin{aligned} f(\boldsymbol{x} \mid \mu, \Sigma) &= \left[P(\mu,\Sigma)\right]^{-T} (2\pi)^{-TN/2} \mid \Sigma \mid^{-T/2} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T}(x_t-\mu)'\Sigma^{-1}(x_t-\mu)\right\} I_R(\boldsymbol{x}) \\ &= \left[P(\mu,\Sigma)\right]^{-T} (2\pi)^{-TN/2} \mid \Sigma \mid^{-T/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(S_x\Sigma^{-1}\right)\right\} I_R(\boldsymbol{x}) \end{aligned} \tag{4}$$

where $S_x = \sum_{t=1}^{T}(x_t-\mu)(x_t-\mu)'$.

As a prior pdf for (μ,Σ), we will use the conventional noninformative diffuse prior (see, for example, Zellner 1971, p.225)

$$f(\mu,\Sigma) \propto \mid \Sigma \mid^{-(N+1)/2} \tag{5}$$

Combining this prior with the likelihood function yields the posterior pdf for (μ,Σ)

$$\begin{aligned} f(\mu,\Sigma \mid \boldsymbol{x}) &\propto f(\boldsymbol{x} \mid \mu, \Sigma)\, f(\mu,\Sigma) \\ &\propto \left[P(\mu,\Sigma)\right]^{-T} \mid \Sigma \mid^{-(T+N+1)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(S_x\Sigma^{-1}\right)\right\} \end{aligned} \tag{6}$$

The presence of the term $P(\mu,\Sigma)$ in this function makes posterior analysis difficult. There is no direct way to integrate out Σ to obtain the marginal posterior pdf for μ, or to integrate out μ to obtain the marginal posterior pdf for Σ. Also, single elements in μ and or Σ are likely to be of interest; there is no direct analytical way of obtaining the marginal posterior pdfs of such single elements. As an alternative, we can sample from these posterior pdfs and use the samples to estimate the marginal posterior pdfs and their moments. The objective of this paper is to describe and illustrate a method for doing so.

2. The Univariate Case

It is convenient to begin by considering the case where x is a truncated univariate normal random variable, truncated from below at a and above at b (a could be $-\infty$ or b could be $+\infty$). Let y be the corresponding non-truncated version of x, with $y \sim N(\mu,\sigma^2)$. In this case equation (3) becomes

$$P(\mu,\sigma) = \Pr(a < y < b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \tag{7}$$

where $\Phi(.)$ is the standard normal cumulative distribution function (cdf). The posterior pdf for (μ,σ^2), equation (6), becomes

$$f(\mu,\sigma^2 \mid \boldsymbol{x}) \propto \left[\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right]^{-T} \left(\sigma^2\right)^{-(T+2)/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{T}(x_t-\mu)^2\right\} \tag{8}$$

It is not possible to analytically integrate out μ or σ^2 from this pdf to obtain the marginal posterior pdfs $f(\sigma^2 \mid \boldsymbol{x})$ and $f(\mu \mid \boldsymbol{x})$. Also, because the conditional posterior pdfs $f(\sigma^2 \mid \mu, \boldsymbol{x})$ and $f(\mu \mid \sigma^2, \boldsymbol{x})$ are not recognisable, it is not possible to set up a Gibbs' sampling algorithm that draws from these conditional pdfs.

Our solution to this problem is to introduce a vector of latent variables $\boldsymbol{y} = (y_1, y_2, \ldots, y_T)'$ that can be viewed as drawings from the non-truncated normal distribution $N(\mu,\sigma^2)$ and that have a direct (deterministic) correspondence with the truncated observations $\boldsymbol{x}$. To appreciate this correspondence, consider the inverse cdf method for drawing observations

y_t from $N(\mu,\sigma^2)$ and observations x_t from $N(\mu,\sigma^2)\times I_{(a,b)}(x_t)$. Given a uniform random draw U from $(0,1)$, draws for y_t and x_t are given respectively by

$$y_t = \mu + \sigma\,\Phi^{-1}\left(U\right) \tag{9}$$

and

$$x_t = \mu + \sigma\,\Phi^{-1}\left[\Phi\left(\frac{a-\mu}{\sigma}\right) + U\left(\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right)\right] \tag{10}$$

The result in equation (9) is well known; the result in equation (10) can be found, for example in Albert and Chib (1996). Equations (9) and (10) can be used for generating a value for the latent variable y_t. Given a value x_t from the truncated distribution, and given (μ,σ^2), we can use equation (10) to compute a value for U

$$U = \frac{\Phi\left(\frac{x_t-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)} \tag{11}$$

Then a corresponding value y_t from the non-truncated distribution can be computed from equation (9)

$$y_t = \mu + \sigma\,\Phi^{-1}\left(U\right) = \mu + \sigma\,\Phi^{-1}\left(\frac{\Phi\left(\frac{x_t-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}\right) \tag{12}$$

We are now in a position to use the values y_t in a Gibbs' sampling algorithm. From Bayes' theorem, we can write the joint posterior pdf for μ,σ^2 and $\boldsymbol{y}$ as

$$\begin{aligned} f(\mu,\sigma^2,\,\boldsymbol{y}\,|\,\boldsymbol{x}) &\propto f(\boldsymbol{x}\,|\,\boldsymbol{y},\mu,\sigma^2)\,f(\boldsymbol{y},\mu,\sigma^2) \\ &= f(\boldsymbol{x}\,|\,\boldsymbol{y},\mu,\sigma^2)\,f(\boldsymbol{y}\,|\,\mu,\sigma^2)\,f(\mu,\sigma^2) \end{aligned} \tag{13}$$

Given the deterministic relationship between $\boldsymbol{x}$ and $\boldsymbol{y}$ defined in equation (12), $f(\boldsymbol{x}\,|\,\boldsymbol{y},\mu,\sigma^2)=1$ when (12) holds, and is zero otherwise. The remaining terms on the right side of (13) involve $\boldsymbol{y}$ not $\boldsymbol{x}$, and so it is possible to express $f(\mu,\sigma^2,\,\boldsymbol{y}\,|\,\boldsymbol{x})$ in terms of the more readily manipulated non-truncated distribution. Specifically,

$$\begin{aligned} f(\mu,\sigma^2,\,\boldsymbol{y}\,|\,\boldsymbol{x}) &\propto \left(\sigma^2\right)^{-(T+2)/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t-\mu)^2\right\} \\ &= \left(\sigma^2\right)^{-(T+2)/2}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{t=1}^{T}(y_t-\bar{y})^2+T(\mu-\bar{y})^2\right]\right\} \end{aligned} \tag{14}$$

where $\bar{y}$ is the sample mean of the y_t and the relationship between $\boldsymbol{y}$ and $\boldsymbol{x}$ is given by (12). The conditional posterior pdfs from equation (14), required for the Gibbs sampler, are

$$f(\mu\,|\,\sigma^2,\boldsymbol{y},\boldsymbol{x}) \propto \exp\left\{-\frac{T}{2\sigma^2}(\mu-\bar{y})^2\right\} \tag{15}$$

$$f(\sigma^2\,|\,\mu,\boldsymbol{y},\boldsymbol{x}) \propto \left(\sigma^2\right)^{-(T+2)/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t-\mu)^2\right\} \tag{16}$$

$$f(\boldsymbol{x}\,|\,\boldsymbol{y},\mu,\sigma^2)=1 \qquad\qquad \text{when (12) holds} \tag{17}$$

These pdfs suggest the following steps for generating (μ,σ^2) from their posterior pdf.

1. Choose starting values for (μ,σ^2).
2. Compute $y_t,\ t=1,2,\ldots,T$ from equation (12).
3. Draw μ from the $N(\bar{y},\ \sigma^2/T)$ distribution in equation (15).
4. Draw σ^2 from the inverted gamma pdf in equation (16).
5. Continue repeating steps 2 to 4, with the conditioning variables being the most recent draws of μ and σ^2, and the most recently calculated values for $\boldsymbol{y}$.

The above procedure is suitable for posterior inferences on the parameters of a univariate truncated normal distribution. To make posterior inferences about the parameters of a multivariate truncated normal distribution we do not employ the above Gibbs' sampler directly, but we build on the results from the univariate case to derive an algorithm for the multivariate case.

4. The Multivariate Case

We return to the posterior pdf for μ and Σ in the multivariate case, namely

$$f(\mu,\Sigma|\,\boldsymbol{x}) \propto \left[P(\mu,\Sigma)\right]^{-T} |\Sigma|^{-(T+N+1)/2} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(S_x\Sigma^{-1}\right)\right\} \tag{18}$$

where $y_t \sim N(\mu,\Sigma)$ and $x_t \sim N(\mu,\Sigma)\times I_R(x_t)$ are now N-dimensional vectors. To use the inverse cdf method to establish a deterministic relationship between $y_t = (y_{1t}, y_{2t},\ldots, y_{Nt})'$ and $x_t = (x_{1t}, x_{2t},\ldots, x_{Nt})'$ we consider a sequence of conditional distributions for the elements in these vectors. Beginning with x_{1t} and y_{1t}, we can write

$$y_{1t} = \tau_1 + \omega_1 \Phi^{-1}\left(\frac{\Phi\left(\frac{x_{1t}-\tau_1}{\omega_1}\right) - \Phi\left(\frac{a_1-\tau_1}{\omega_1}\right)}{\Phi\left(\frac{b_1-\tau_1}{\omega_1}\right) - \Phi\left(\frac{a_1-\tau_1}{\omega_1}\right)}\right) \tag{19}$$

where $\omega_1 = \sqrt{\sigma_{11}}$ is the square root of the first diagonal element in Σ and $\tau_1 = \mu_1$.

To compute a value for y_{2t} we consider the distribution of y_{2t} conditional on y_{1t}. This distribution has mean and standard deviation given by

$$E(y_{2t}\,|\,y_{1t}) = \mu_2 + \sigma_{12}\sigma_{11}^{-1}(y_{1t}-\mu_1) = \tau_{2t} \tag{20}$$

$$\operatorname{sd}(y_{2t}\,|\,y_{1t}) = \left(\sigma_{22} - \sigma_{12}\sigma_{11}^{-1}\sigma_{21}\right)^{1/2} = \omega_2 \tag{21}$$

where σ_{ij} is the (i, j)-th element of Σ and μ_i is the i-th element of μ. The value for y_{2t} can be calculated from

$$y_{2t} = \tau_{2t} + \omega_2 \Phi^{-1}\left(\frac{\Phi\left(\frac{x_{2t} - \tau_{2t}}{\omega_2}\right) - \Phi\left(\frac{a_2 - \tau_{2t}}{\omega_2}\right)}{\Phi\left(\frac{b_2 - \tau_{2t}}{\omega_2}\right) - \Phi\left(\frac{a_2 - \tau_{2t}}{\omega_2}\right)} \right) \tag{22}$$

We can continue in this way considering the distribution of $(y_{3t} | y_{1t}, y_{2t})$, then $(y_{4t} | y_{1t}, y_{2t}, y_{3t})$ and so on. Expressions for the conditional means and standard deviations can be found, for example, in Judge et al (1988, p.50). Those for $(y_{3t} | y_{1t}, y_{2t})$ are

$$E(y_{3t} | y_{1t}, y_{2t}) = \tau_{3t} = \mu_3 + \begin{pmatrix} \sigma_{31} & \sigma_{32} \end{pmatrix} \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}^{-1} \begin{pmatrix} y_{1t} - \mu_1 \\ y_{2t} - \mu_2 \end{pmatrix} \tag{23}$$

$$\mathrm{sd}(y_{3t} | y_{1t}, y_{2t}) = \omega_3 = \left[\sigma_{33} - \begin{pmatrix} \sigma_{31} & \sigma_{32} \end{pmatrix} \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{13} \\ \sigma_{23} \end{pmatrix} \right]^{1/2} \tag{24}$$

The generalization to $i = 4, 5, \ldots,$ etc is straightforward.

Proceeding in this way for all sample observations, establishes a relationship between $\boldsymbol{x} = (x_1, x_2, \ldots, x_T)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_T)'$. Analogous to equation (13), we can write Bayes theorem as

$$\begin{aligned} f(\mu, \Sigma, \boldsymbol{y} | \boldsymbol{x}) &\propto f(\boldsymbol{x} | \boldsymbol{y}, \mu, \Sigma)\ f(\boldsymbol{y}, \mu, \Sigma) \\ &= f(\boldsymbol{x} | \boldsymbol{y}, \mu, \Sigma)\ f(\boldsymbol{y} | \mu, \Sigma)\ f(\mu, \Sigma) \end{aligned} \tag{25}$$

The pdf $f(\boldsymbol{x} | \boldsymbol{y}, \mu, \Sigma)$ is equal to one with the exact relationship between $\boldsymbol{x}$ and $\boldsymbol{y}$ being defined by equations (19) and (22) and their extensions to the later elements in x_t and y_t. Then, the posterior pdf for μ and Σ, written in terms of the y_t, is

$$f(\mu,\Sigma,\boldsymbol{y}\,|\,\boldsymbol{x}) \propto |\Sigma|^{-(T+N+1)/2} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(S_y\Sigma^{-1}\right)\right\} \tag{26}$$

where $S_y = \sum_{t=1}^{T}(y_t-\mu)(y_t-\mu)'$.

For a Gibbs' sampling algorithm we need the conditional posterior pdfs from (26). The conditional posterior pdf for Σ is the inverted Wishart pdf

$$f(\Sigma\,|\,\mu,\boldsymbol{y},\boldsymbol{x}) \propto |\Sigma|^{-(T+N+1)/2} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(S_y\Sigma^{-1}\right)\right\} \tag{27}$$

To establish the conditional posterior pdf for μ, note that

$$S_y = \sum_{t=1}^{T}\left(y_t-\bar{y}\right)\left(y_t-\bar{y}\right)' + T\left(\mu-\bar{y}\right)\left(\mu-\bar{y}\right)' \tag{28}$$

where $\bar{y}$ is the sample mean of the y_t. Also,

$$\operatorname{tr}\left[T\left(\mu-\bar{y}\right)\left(\mu-\bar{y}\right)'\Sigma^{-1}\right] = \left(\mu-\bar{y}\right)'\left(\frac{\Sigma}{T}\right)^{-1}\left(\mu-\bar{y}\right) \tag{29}$$

Using equations (28) and (29) in equation (26), we can establish that the conditional posterior pdf for μ is the multivariate normal distribution

$$f\left(\mu\,|\,\Sigma,\boldsymbol{y},\boldsymbol{x}\right) \propto \exp\left\{-\frac{1}{2}\left(\mu-\bar{y}\right)'\left(\frac{\Sigma}{T}\right)^{-1}\left(\mu-\bar{y}\right)\right\} \tag{30}$$

We are now in a position to summarize the Gibbs' sampling procedure for drawing observations (μ,Σ) from their posterior pdf.

1. Choose starting values for (μ,Σ).

2. Compute y_{it}, $i = 1,2,\ldots,N$ and $t = 1,2,\ldots,T$ using the expressions in (19) and (22) and their generalizations, and using the values for τ_{it} and ω_i defined below equation (19) and in equations (20), (21), (23), (24) and their extensions.
3. Draw μ from the $N(\bar{y}, \Sigma/T)$ distribution in (30).
4. Draw Σ from the inverted Wishart distribution in equation (27).
5. Continue repeating steps 2 to 4, with the conditioning variables being the most recent draws of μ and Σ, and the most recently calculated values of $\boldsymbol{y}$.

5. The Application

The variable chosen for an example is rainfall over the four months from January to April in five shires in the northern part of the Western Australian wheat belt: Northampton, Chapman Valley, Mullewa, Greenough and Irwin. Rainfall data were obtained from the Western Australian office of the Bureau of Meteorology as part of another study concerned with predictive densities for shire-level wheat yield (Griffiths et al 2001). However, that study used rainfall data over the months May to October, not the first four months of the year that we are considering here. The vector x_t is of dimension (5×1) containing the four-month rainfalls for each of the five shires in year t. There are 49 observations ranging from 1950 to 1998. The rainfall for a given shire was taken as the measured rainfall at a site considered representative of that shire. These sites were Northampton P.O. (for Northampton shire), Chapman Research Station at Nabawa (for Chapman Valley shire), Mullewa (for Mullewa shire), Geraldton airport (for Greenough shire), and Dongara (for Irwin shire). Each rainfall distribution is assumed to be truncated from below at zero and not truncated from above. Thus, we have $a_i = 0$ and $b_i = \infty$ for $i = 1,2,\ldots,5$. The Gibbs' sampler was used to generate a total of 12,000 observations with the first 2,000 discarded as a burn in. Plots of the generated observations showed no evidence of nonstationarity.

Histograms and summary statistics for the rainfall data are graphed in Figure 1. The unit of measurement is millimetres. The rainfall distributions tend to be concentrated between zero and 50 millimetres and then tail off to the right. In shires 2 and 3 there is some evidence of a second mode, around 90 and 140 millimetres, respectively. The bimodality could be attributed to the fineness of the histogram, however. It seems unlikely that bimodality would persist if a larger sample was taken, and so we proceed with the truncated normal distribution assumption. The marginal posterior pdfs for the parameters μ_i for each shire, and summary statistics for these pdfs, appear in Figure 2, adjacent to the rainfall graphs for each shire. In all shires except the first, the posterior pdf for each μ_i is approximately symmetric and, as one would expect for a truncated distribution, centred around a value to the left of the sample mean. The posterior pdf for μ_1 is skewed to the left and has a mean of -44.76. Ignoring the effect of correlations with other shires, this outcome suggests a mode at zero and that rainfall is modeled via the right tail of a normal distribution.

As an example of the posterior pdfs for some of the elements in Σ, those for $\sqrt{\sigma_{33}}$ and $\sqrt{\sigma_{44}}$, and related summary statistics, appear in Figure 3. These pdfs are skewed to the right and centred around values higher then the sample standard deviations of the truncated distributions. Finally, to give an idea of the correlation between rainfalls of adjacent shires the posterior pdfs for $\rho_{12} = \sigma_{12} \big/ \sqrt{\sigma_{11}\sigma_{22}}$ and $\rho_{34} = \sigma_{34} \big/ \sqrt{\sigma_{33}\sigma_{44}}$ are presented in Figure 4. These pdfs are skewed to the left and with means of 0.89 and 0.72, they suggest high correlations between the rainfalls.

6. Concluding Remarks

We have demonstrated how a relatively simple Gibbs' sampler can be set up to find posterior pdfs for the parameters of a truncated multivariate normal distribution. In the rainfall example this information could be utilized further to obtain predictive pdfs for rainfall. These predictive pdfs can then be used to incorporate rainfall uncertainty into predictive pdfs for wheat yield or into other models with outcomes that depend on rainfall. The algorithm is potentially useful in other areas where truncated distributions are utilized, such as in the area of stochastic frontier production functions.

References

Albert, J.H. and Chib, S. (1996), “Computation in Bayesian Econometrics: An Introduction to Markov Chain Monte Carlo”, in Hill, R.C. (ed.), *Advances in Econometrics Volume 11A: Computational Methods and Applications*. JAI Press, Greenwich, pp.3-24.

Greene, W.H. (1997), “Frontier Production Functions”, in Pesaran, M.H. and Schmidt, P. (eds.), *Handbook of Applied econometrics, Volume II: Microeconometrics*. Blackwell, Malden, pp.81-166.

Griffiths, W.E., Newton, L.S. and O’Donnell, C.J. (2001), “Predictive Densities for Shire Level Wheat Yield In Western Australia”, Paper contributed to the Australian Agricultural and Resource Economics Society Conference, Adelaide.

Judge, G.G., Hill, R.C., Griffiths, W.E., Lütkepohl, H., Lee, T.-C. (1988), *Introduction to the Theory and Practice of Econometrics*, 2nd edition. John Wiley and Sons, New York.

Koop, G. and M.F.J. Steel (2001), “Bayesian Analysis of Stochastic Frontier Models”, in Baltagi, B.H. (ed.), *A Companion to Theoretical Econometrics*, Blackwell, Oxford.

Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*. John Wiley and Sons, New York.

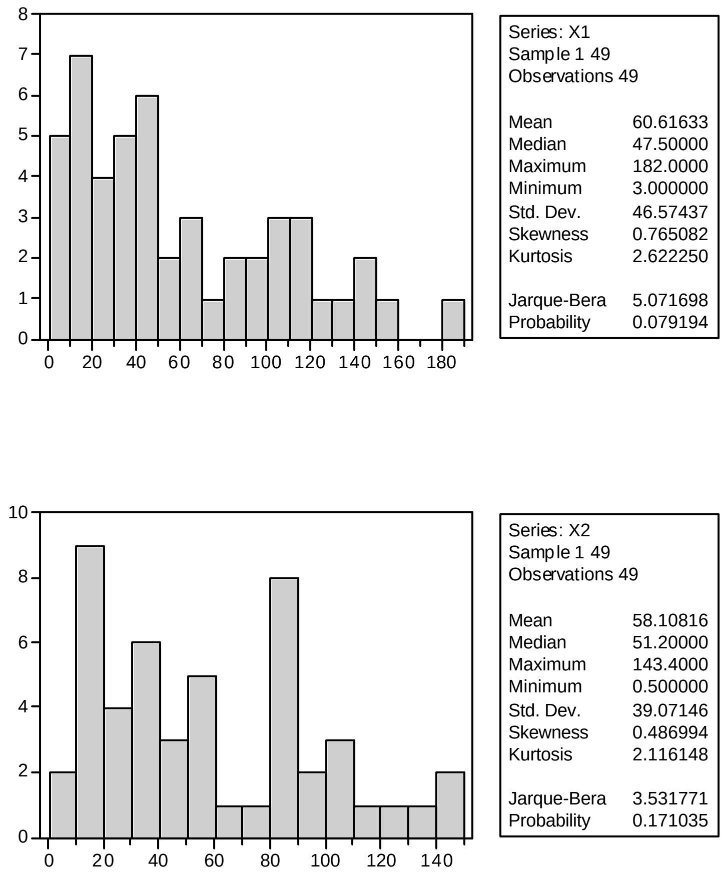

Figure 1(a) Histograms and summary statistics for rainfall data in Shires 1 and 2

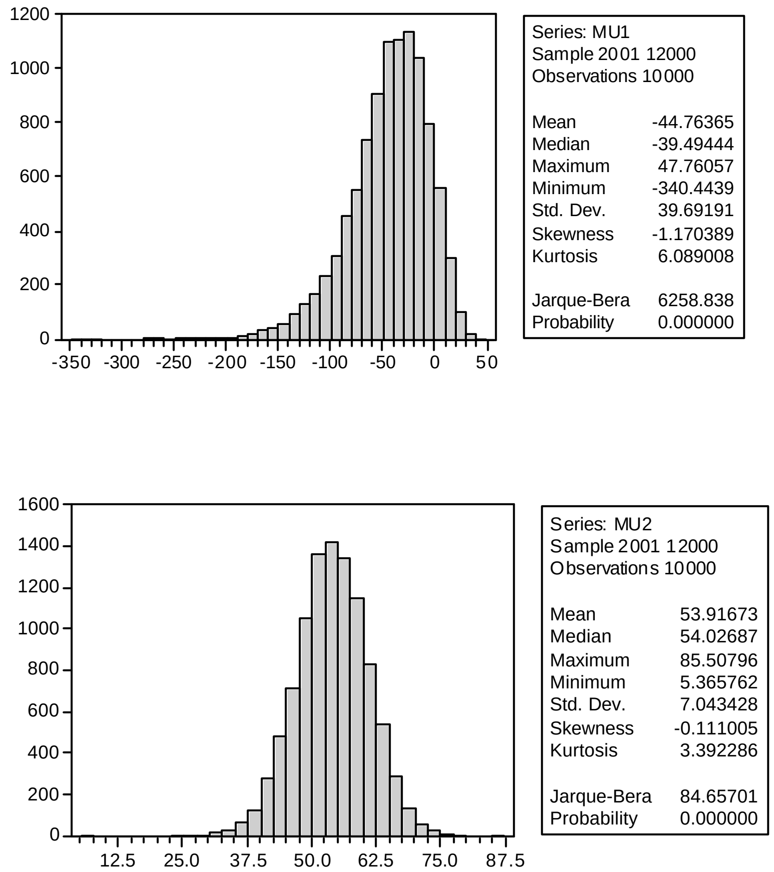

Figure 2(a) Posterior pdfs and summary statistics for μ_1, μ_2

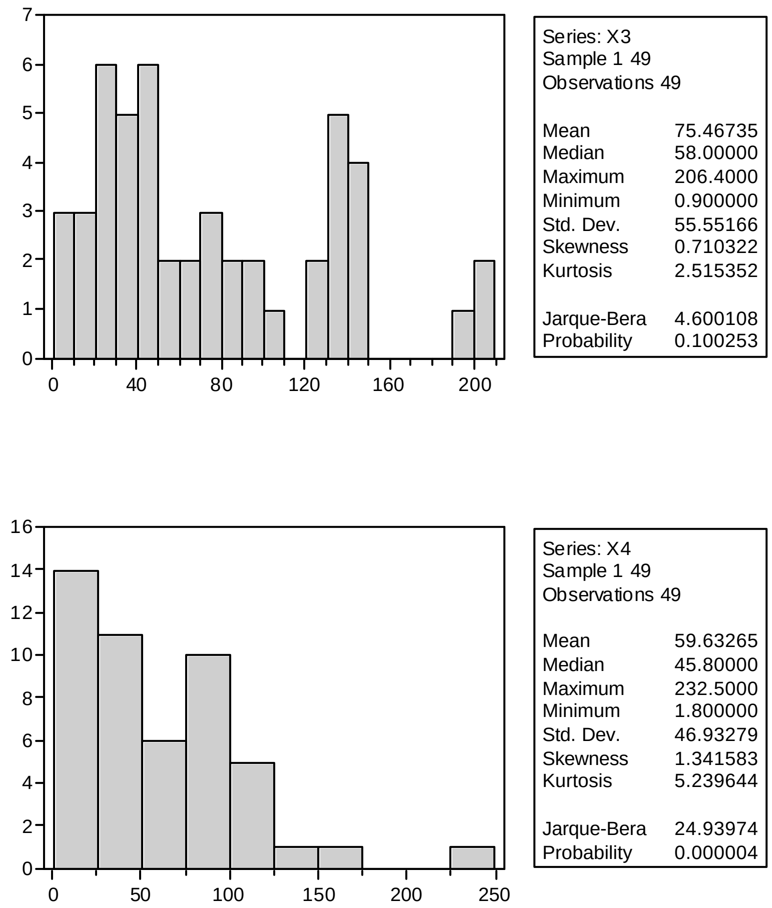

Figure 1(b) Histograms and summary statistics for rainfall data in Shires 3 and 4

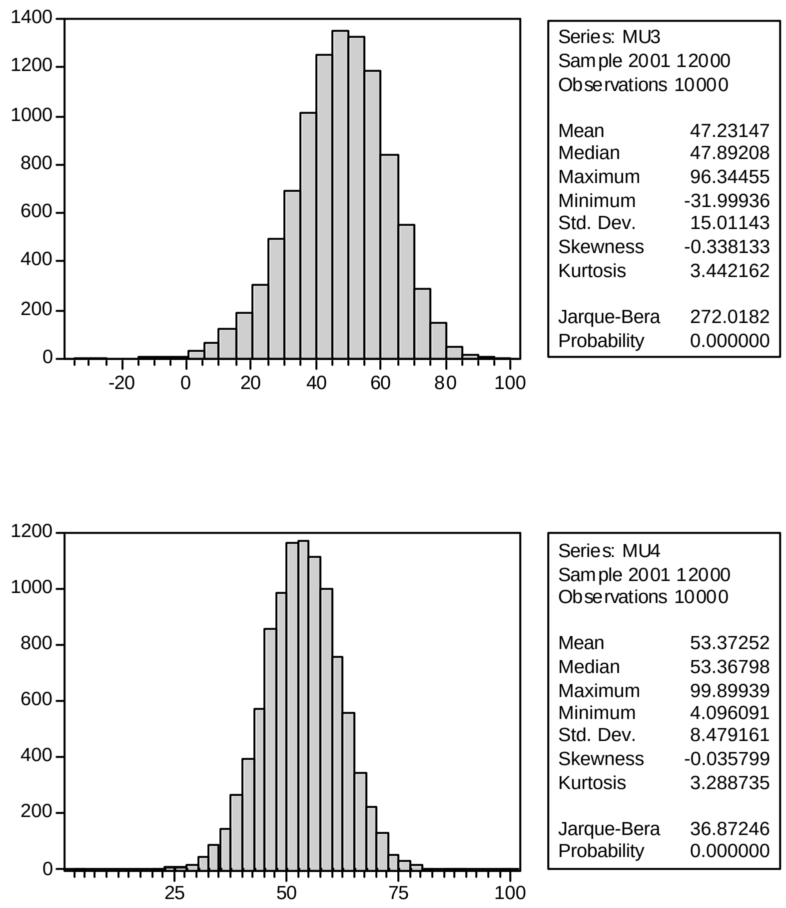

Figure 2(b) Posterior pdfs and summary statistics for μ_3, μ_4

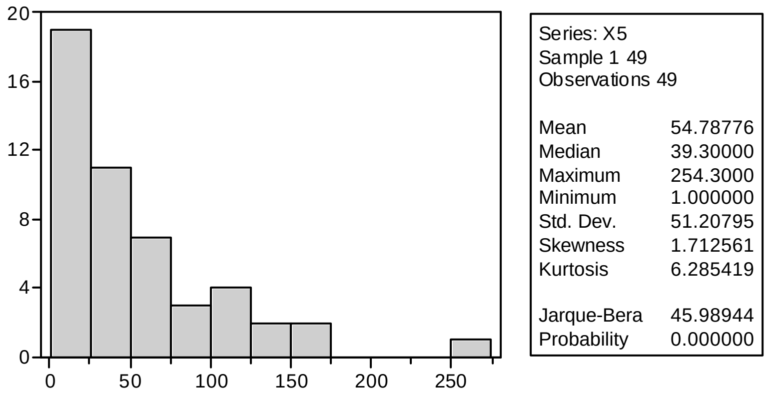

Figure 1(c) Histograms and summary statistics for rainfall data in Shire 5

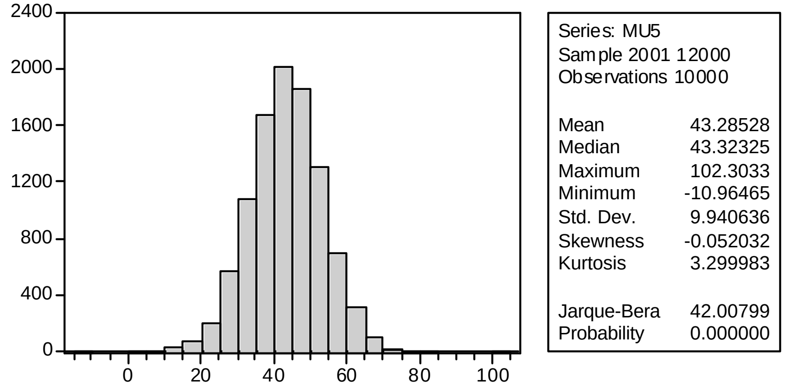

Figure 2(c) Posterior pdfs and summary statistics for μ_5

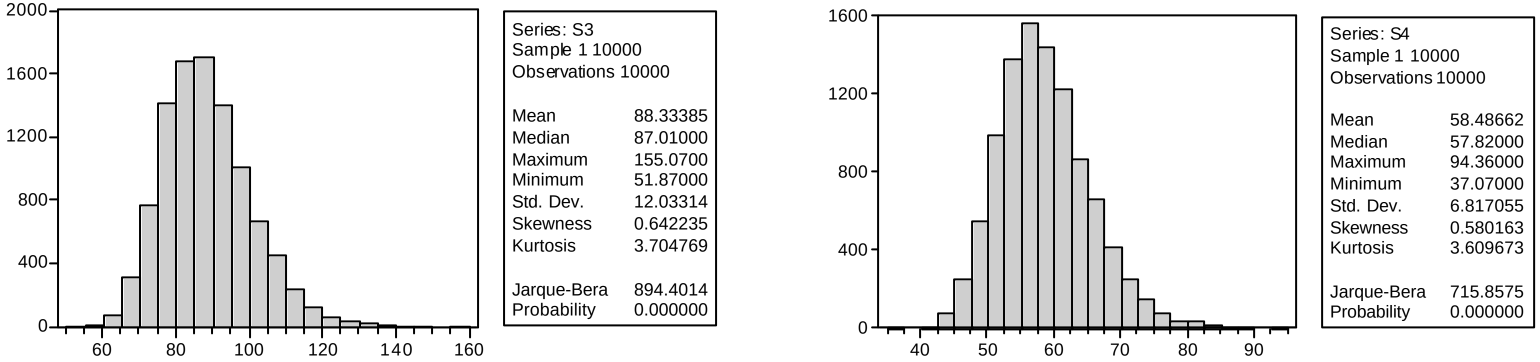

Figure 3 Posterior pdfs for $\sqrt{\sigma_{33}}$ and $\sqrt{\sigma_{44}}$

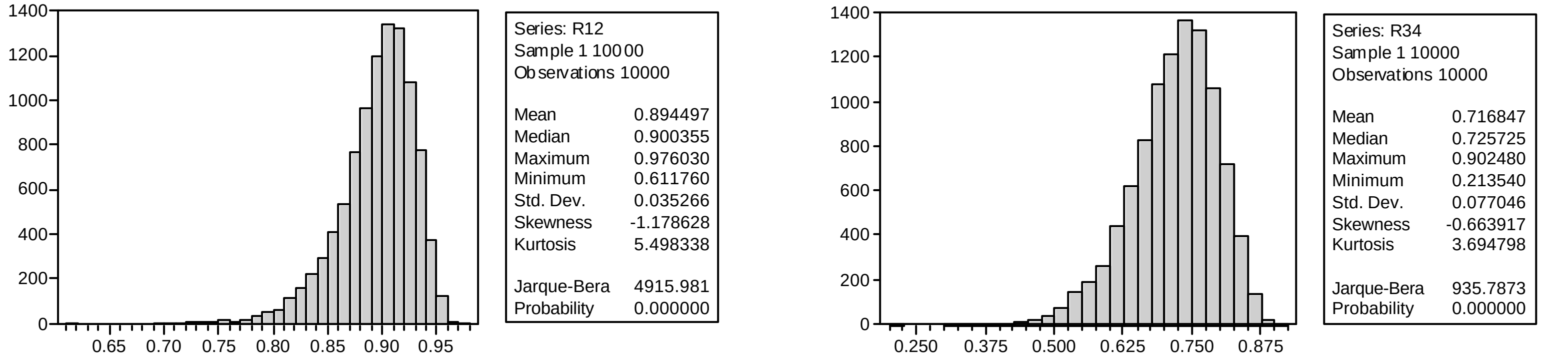

Figure 4 Posterior pdfs for ρ_{12} and ρ_{34}